Chemistry & Biology
Article

Prediction and Manipulation of the Stereochemistry of Enoylreduction in Modular Polyketide Synthases

David H. Kwan,[1,2] Yuhui Sun,[1,2] Frank Schulz,[1] Hui Hong,[1] Bojana Popovic,[1] Joalice C.C. Sim-Stark,[1] Stephen F. Haydock,[1] and Peter F. Leadlay[1,*]
[1]Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK
[2]These authors contributed equally to this work.
*Correspondence: pfl10@mole.bio.cam.ac.uk
DOI 10.1016/j.chembiol.2008.09.012

SUMMARY

When an enoylreductase enzyme of a modular polyketide synthase reduces a propionate extender unit that has been newly added to the growing polyketide chain, the resulting methyl branch may have either *S* or *R* configuration. We have uncovered a correlation between the presence or absence of a unique tyrosine residue in the ER active site and the chirality of the methyl branch that is introduced. When this position in the active site is occupied by a tyrosine residue, the methyl branch has *S* configuration, otherwise it has *R* configuration. In a model PKS in vivo, a mutation (Tyr to Val) in an erythromycin PKS-derived ER caused a switch in the methyl branch configuration in the product from *S* to *R*. In contrast, alteration (Val to Tyr) at this position in a rapamycin-derived PKS ER was insufficient to achieve a switch from *R* to *S*, showing that additional residues also participate in stereocontrol of enoylreduction.

INTRODUCTION

Complex or reduced polyketides constitute one of the largest and most structurally diverse classes of natural product, including numerous clinically valuable antibacterial, antifungal, antitumor, and immunosuppressant compounds. They are produced, most prolifically by *Streptomyces* and allied filamentous bacteria, by stepwise construction of the polyketide chain on modular polyketide synthases (PKSs), giant assembly line multienzymes in which each successive set (or module) of fatty acid synthase-related enzymes normally catalyzes the extension of the polyketide chain by one unit, the intermediates remaining tethered to the multienzyme during assembly (Staunton and Weissman, 2001). Many striking mechanistic and structural similarities have been observed between bacterial modular PKSs and mammalian fatty acid synthases, but PKSs generate much greater structural diversity in their products (Smith and Tsai, 2007). A given PKS may recruit one of a variety of starter units, activated as their CoA esters; in addition, different extender units may be used in each module (usually either acetate units from malonyl-CoA or propionate units from (2*S*)-methylmalonyl-CoA). Furthermore, the normal reductive activities of a fatty acid synthase (ketoreductase [KR], dehydratase [DH], and enoylreductase [ER] activities), which normally accomplish together the full reduction of the initially formed β-ketoacyl group on the nascent chain, may or may not be present in each PKS module, leading to different oxidation states and a large number of alternative chemical structures (Staunton and Weissman, 2001; Smith and Tsai, 2007).

The PKS-catalyzed assembly process also generates stereochemical diversity, because carbon-carbon double bonds may have either *cis* or *trans* geometry, and because of the chirality of centers bearing hydroxyl groups (where they are retained) and branching methyl groups (the latter arising from use of propionate extender units). Understanding how PKSs exert control over the configuration of these stereocenters is crucial in rationalizing longstanding observations, codified as Celmer's rules (Celmer, 1965), of position-specific homology between natural macrocyclic polyketides, as well as for current attempts to create novel polyketide products by manipulating PKS genes. Key elements of the molecular basis of this stereocontrol have already been established. For example, if methyl-bearing centers in the polyketide chain are found to have opposite configurations, even though only the (2*S*)-isomer of methylmalonyl-CoA is used as a substrate (Marsden et al., 1994; Wiesmann et al., 1995), it is in some cases because, in certain modules, after condensation catalyzed by the KS domain, epimerization of the initially formed (2*R*)-methyl-3-ketoacyl thioester allows access to the opposite configuration at the C-2 methyl branch (Weissman et al., 1997).

The chirality of C-3 hydroxyl groups is dictated by the KR domain, and comparison of amino acid sequences of KR domains where the stereochemical outcome is known, together with homology modeling, has allowed the identification of specific amino acid residues at the KR active site in modular PKSs that appear to correlate with the stereochemical outcome (Caffrey, 2003; Reid et al., 2003; Baerga-Ortiz et al., 2006; O'Hare et al., 2006). From these experiments, it is clear that small energetic differences arising from precisely placed "gatekeeper" residues at the KR active site lead to the presentation of a different face of the polyketide substrate to an otherwise essentially identical catalytic site. The presence of these sequence motifs allows confident prediction of the configuration of the polyketide product, a potentially valuable tool for synthetic chemists wishing to confirm the configuration of a target natural product for which the biosynthetic genes are known (Janssen et al., 2007; Udwary et al., 2007; Bock et al., 2008). A more elaborate extension of these rules for ketoreduction has recently been proposed (Keatinge-Clay, 2007) based on X-ray crystal structures of two PKS KR domains.

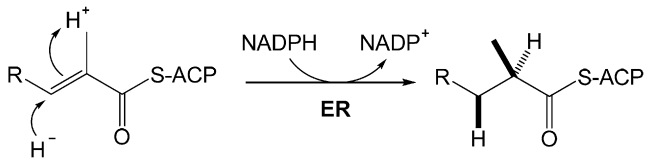

Figure 1. The Reaction Catalyzed by ER Domains
1,4-Nucleophilic addition of hydride ion, delivered from NADPH, to the unsaturated thioester is followed by stereospecific protonation, which establishes the configuration of the methyl branch. *Anti* addition to the double bond is illustrated, as demonstrated for mammalian fatty acid synthase (Smith and Tsai, 2007), but instances of *syn* addition are also known (Smith and Tsai, 2007).

In contrast, it has previously been impossible to predict the stereochemical outcome of enoylreduction by the ER domain, or to account for it by features of the active site. We now report an unexpectedly precise correlation between a unique amino acid residue in an ER domain and the chirality of the methyl branch produced when a propionate extender unit is fully reduced in a PKS module by the action of that ER. Furthermore, homology modeling has revealed that the diagnostic residue is in the active site, and specific mutagenesis has confirmed its importance for stereocontrol. The correlation holds true not only for antibiotic-producing PKS multienzymes, but also for similar PKSs that control the production of long-chain methyl-branched fatty acids forming the highly complex cell walls of *Mycobacterium tuberculosis* and other disease-causing mycobacteria (Gokhale et al., 2007).

RESULTS

As shown in Figure 1, the accepted mechanism of ER-catalyzed reduction, deduced from studies on fatty acid synthases (Smith and Tsai, 2007), involves 1,4-nucleophilic addition of a hydride ion from the coenzyme NADPH to the unsaturated thioester intermediate, followed by stereospecific protonation at the α-carbon, which establishes the configuration of the methyl branch. Our initial hypothesis was that all PKS ER domains have essentially the same active site architecture, and act upon *trans*-alkene substrates by the same mechanism. Opposite configurational outcomes are dictated by relatively minor changes in the presentation of the substrate to the active site as a result of differences in a few critical amino acid residues, leading to protonation from the opposite face of the double bond. For the vast majority of ER domains, the geometry of the substrate (an unsaturated PKS-bound intermediate that does not accumulate) has not been established by experiment. However, where this has been done, for PikER4 (Castonguay et al., 2007) and for erythromycin PKS ER4 (EryER4) (Kellenberger et al., 2008), the geometry was confirmed as *trans*. We were therefore encouraged to examine publicly available ER sequences for residues that might be diagnostic of the stereochemical outcome of a given ER domain.

Multiple Sequence Alignment of ER Domains Reveals a Unique Amino Acid Residue Predictive of the Stereochemical Outcome of Enoylreduction

Multiple sequence alignments were constructed for all available type I ER domains, in which a propionate unit becomes fully reduced and gives rise to a methyl branch of known configuration in the final polyketide product. Scrutiny of these sequence alignments (Figure 2; see Supplemental Data available online) revealed that there is a single amino acid that shows excellent correlation with the configuration of the polyketide product for those cases where it is known. This residue lies about 90 residues N terminal of a well-conserved HAAAGGVGMA consensus sequence identified as the site of NADPH cofactor binding (Amy et al., 1989). In those ER domains producing a (2*S*)-methyl branch, a tyrosine (Y) residue is systematically conserved at this position, whereas in those ER domains producing a (2*R*)-methyl branch, a valine (V) residue (or occasionally alanine or phenylalanine) is found at this position (Figure 2). This position corresponds precisely to Tyr52 in the closely homologous *Escherichia coli* enzyme quinone oxidoreductase (QOR; PDB ID 1QOR) (Thorn et al., 1995). For ease of reference, this position in the PKS ER domains is labeled 52′, and other residues within the ER domain are numbered relative to it. Among the ER sequences analyzed, those processing a propionate extender unit all appear to follow this trend, the only exception (Supplemental Data) appearing to be an ER domain from the myxobacterial soraphen PKS (Bedorf et al., 1993), soraphen ER3, which acts on an unusual methoxy-branched extender unit (Wenzel et al., 2006). The product of this PKS has been confirmed by X-ray crystallography to have *S* configuration at the relevant position, but the ER domain has a leucine instead of a tyrosine at residue 52′ (Ligon et al., 2002).

Homology Models of PKS ER Domains Place Residue 52′ at the Active Site

As previously reported (Maier et al., 2006), searches of publicly available databases readily show that the ER domains of modular PKSs are members of the medium-chain NAD(P)H-dependent dehydrogenase/reductase (MDR) family of enzymes, and that they share most significant homology with the QOR subfamily (Persson et al., 1994). Homology models of the ERs from 6-deoxyerythronolide B synthase (DEBS) module 4 (EryER4) and rapamycin synthase (RAPS) module 13 (rapamycin PKS ER13 [RapER13]) were constructed based on the MDR enzyme *E. coli* QOR (PDB ID 1QOR) (Thorn et al., 1995) as template (24% sequence identity), as described in the Experimental Procedures. The overall fold predicted for both EryER4 and RapER13 was essentially identical, and served to reveal key conserved features of their predicted active sites (Figure 3). The position of the bound NADPH cofactor could also be modeled with confidence. Both model structures suggest that residue 52′ (Tyr in EryER4; Val in RapER13) lies at the active site, adjacent to the C4′ of NADPH, from which the hydride ion is transferred during reduction (Figure 3).

Site-Directed Mutagenesis of ER Domains in a Triketide Synthase Multienzyme

The bimodular PKS DEBS1-TE (Cortés et al., 1995), housed in a suitable strain of *Saccharopolyspora erythraea* or of a related actinomycete, is an established model system for the study of modular polyketide synthesis. It was engineered by relocating the chain-terminating thioesterase domain from DEBS3 to the C terminus of the bimodular DEBS1 protein, and the products of this synthase are triketide lactones. In previous work, it has

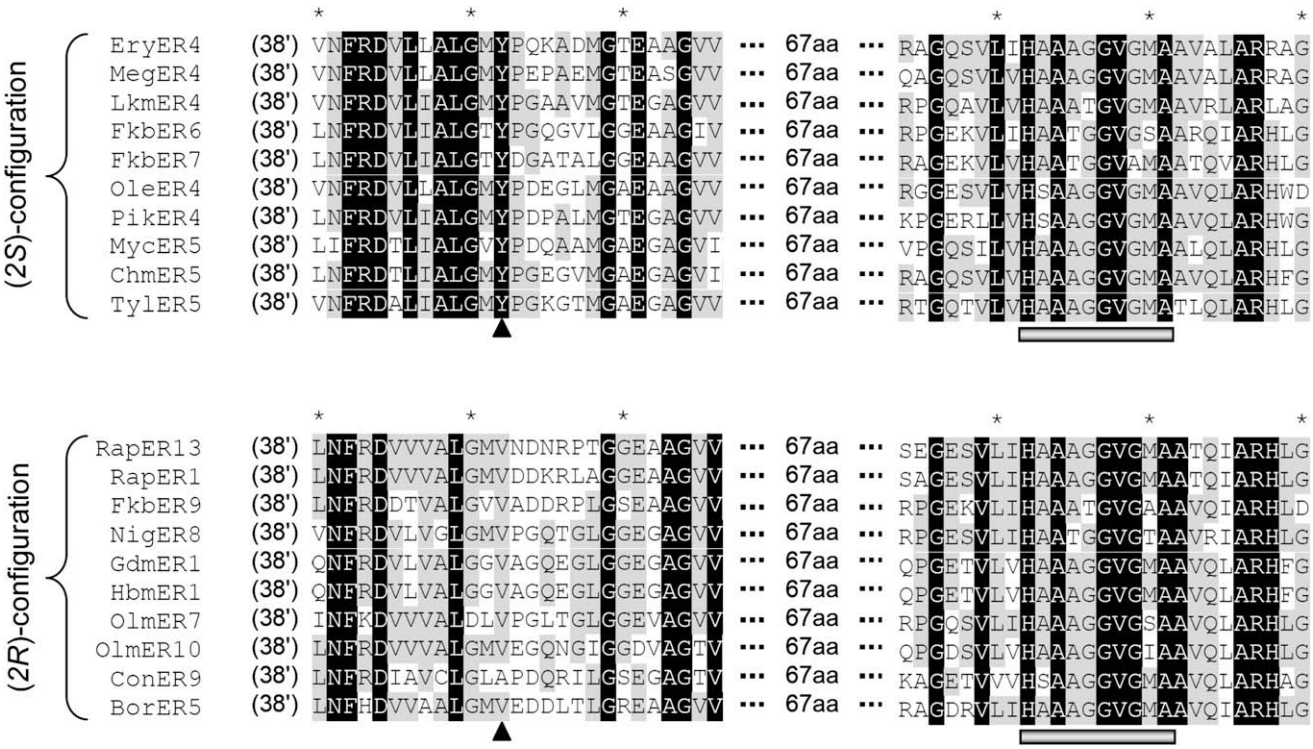

Figure 2. Sequence Alignment of a Portion of Representative ER Domains from PKSs

The position of the unique tyrosine (Y52′) residue correlated with (2*S*) configuration in the polyketide product is marked with black triangles. The position of the NADPH binding site is marked with gray bars. Residue numbering is that of *E. coli* QOR (PDB ID 1QOR) (Thorn et al., 1995) on which the ER domains were modeled.

been shown that, when the KR2 domain of DEBS1-TE is replaced by the full "reductive loop" (Kellenberger et al., 2008) of KR, DH, and ER domains from DEBS module 4, the resulting triketide synthase, referred to here as TKS-ery4, produces the triketide lactones **1a** and **1b** fully reduced at C-3 (the **a** and **b** forms result from the incorporation of either propionate or acetate as starter units, respectively) (Figure 4A). In contrast, replacement of the module 2 KR domain of DEBS1-TE by the reductive loop (KR, DH, and ER domains) of RAPS module 13 resulted in a triketide synthase, TKS-rap13, which produced triketide lactones **2a** and **2b** (Figure 4B) (Kellenberger et al., 2008).The compounds **1a** and **1b** differ from compounds **2a** and **2b** in the configuration of the methyl substituent at the C-2 position, a difference that arises presumably from the differing stereospecificity of the EryER4 and the RapER13 domains (Kellenberger et al., 2008). The products of TKS-ery4 have (2*S*) configuration consistent

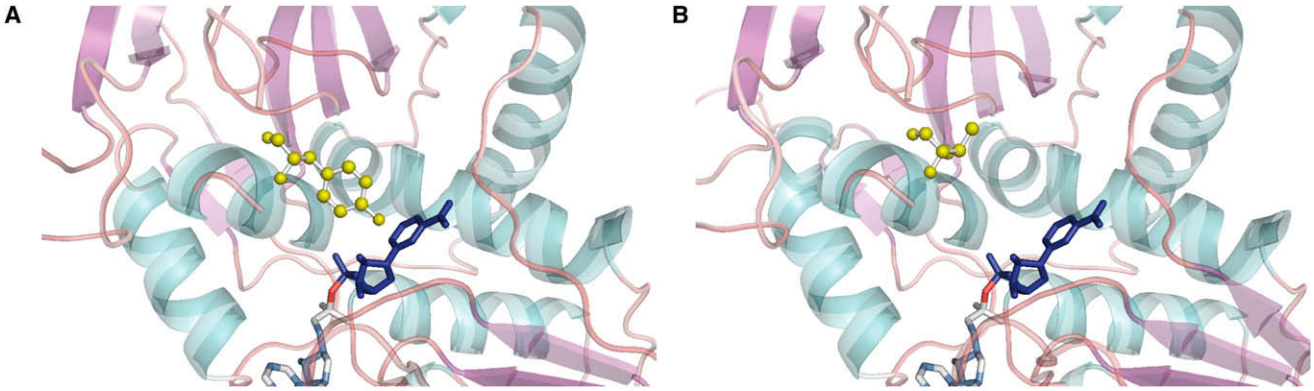

Figure 3. Homology Models of ER Domains

(A and B) Models of (A) EryER4 and (B) RapER13 with bound cofactor NADPH were constructed using *E. coli* QOR (PDB ID 1QOR) (Thorn et al., 1995) as template. The characteristic tyrosine (EryER4) and valine (RapER13) residues are highlighted in yellow, and NADPH is shown in dark blue.

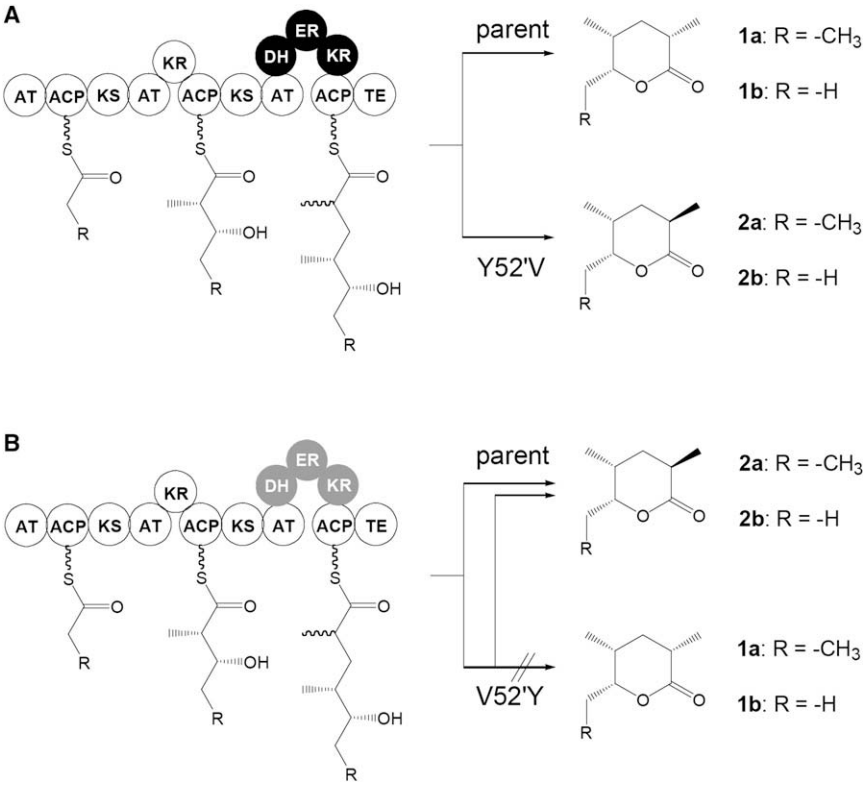

Figure 4. Recombinant Triketide Synthases Based on DEBS1-TE

(A and B) Schematic representation of the DEBS1-TE-based triketide synthase (Cortés et al., 1995) with (A) the reductive loop (Kellenberger et al., 2008) of *ery* module 4, or (B) the reductive loop (Kellenberger et al., 2008) of *rap* module 13. The products of the parent enzyme are shown alongside those of the variants mutated at residue 52′.

with having Tyr at position 52′, while the products of TKS-rap13 have (2*R*) configuration consistent with having Val at position 52′.

To analyze the role of residue 52′ more precisely, site-directed mutagenesis was carried out at this position within the ER domains of TKS-ery4 (Y52′V) and TKS-rap13 (V52′Y) (Figure 4). The mutant DEBS1-TE genes were cloned (see Experimental Procedures) into the integrative plasmids pYH141 and pYH137, under the control of the inducible *actI-actIIorf4* promoter (McDaniel et al., 1994), and introduced by conjugation into *Streptomyces coelicolor* CH999 (McDaniel et al., 1994), where they integrated into the genome. Cultures of the recombinant strains produced low (~1–2 mg/L), but readily detectable levels of triketide lactones that were not produced by the plasmid-free control strain. In this host, the Y52′V TKS-ery4 mutant produced compounds **2a** and **2b** as judged by LC-MS analysis (Supplemental Data) and not the compounds **1a** and **1b** produced by the strain housing the parent PKS, TKS-ery4. This change represents a switch in the configuration of the C-2 methyl group. In contrast, the V52′Y TKS-rap13 mutant produced the same compounds **2a** and **2b** as the strain housing the parent PKS, TKS-rap13 (Supplemental Data). Authentic synthetic samples of the known compounds **1a**, **1b**, **2a**, and **2b**, used for calibration of the LC-MS, were the kind gift of Dr. M.L. Heathcote (Department of Chemistry, University of Cambridge).

These initial results prompted further investigation using as the host strain *S. erythraea* BIOT1717-JC2, an erythromycin-overproducing strain from which essentially all erythromycin PKS genes had been deleted. The PKS genes encoding TKS-ery4 and TKS-rap13, as well as the mutant genes encoding Y52′V TKS-ery4 and V52′Y TKS-rap13, were each introduced into BIOT1717-JC2 under the control of the erythromycin PKS promoter, on suitable integrative plasmids. Fermentation of the resulting recombinant strains gave significantly increased yields of triketide lactones, all of which were **a** forms (propionate starter unit derived), consistent with the fact that the BIOT1717 strain had originally been optimized for erythromycin A production, which requires exclusively propionate starter units. In agreement with the observations made using *S. coelicolor*, LC-MS analysis of broth extracts showed that the strain encoding TKS-ery4 produced chiefly **1a**, but in this strain it was accompanied by small amounts of **2a**, which would not have been detectable in the lower-yielding *S. coelicolor* strain (Figure 5A). The BIOT1717-JC2 strain containing Y52′V TKS-ery4 produced product **2a** only, in amounts similar to that of the strain carrying the parent TKS-ery4 (Figure 5B). The strain encoding TKS-rap13, and the strain encoding V52′Y TKS-rap13, both produced **2a** at nearly equal levels, and with no detectable amount of the other isomer (Figures 5C and 5D). The products of strains carrying TKS-ery4 and Y52′V TKS-ery4 were coinjected into the LC-MS and confirmed to migrate as separate peaks.

Prediction of the Configuration of Methyl-Branched Cell Wall Lipids in Mycobacteria

Within their complex cell walls, the pathogenic mycobacteria contain multiply methyl-branched, saturated lipids of polyketide or fatty acid origin. These cell wall lipids give rise to the waxy envelope that is characteristic of mycobacteria, and are considered to be a major determinant of mycobacterial pathogenicity, with important roles in the host-immune response (Gokhale et al., 2007). We were interested to compare the experimentally determined methyl-branch configuration of these lipids to the configuration predicted on the basis of the indicator residue that we had identified in the ER domains of the very similar PKS enzymes of antibiotic-producing actinomycetes. The dimycocerosate esters (DIMs), for example, are mycobacterial cell wall lipids that contribute to the permeability barrier of the cell envelope and have been shown to promote virulence in pathogenic mycobacteria (Cox et al., 1999; Camacho et al., 2001). The DIMs consist of two multiply methyl-branched, saturated acyl chains, mycocerosic acids, attached in ester linkage to neighboring hydroxyl groups in a phthiocerol moiety (Figure 6). Both mycocerosyl

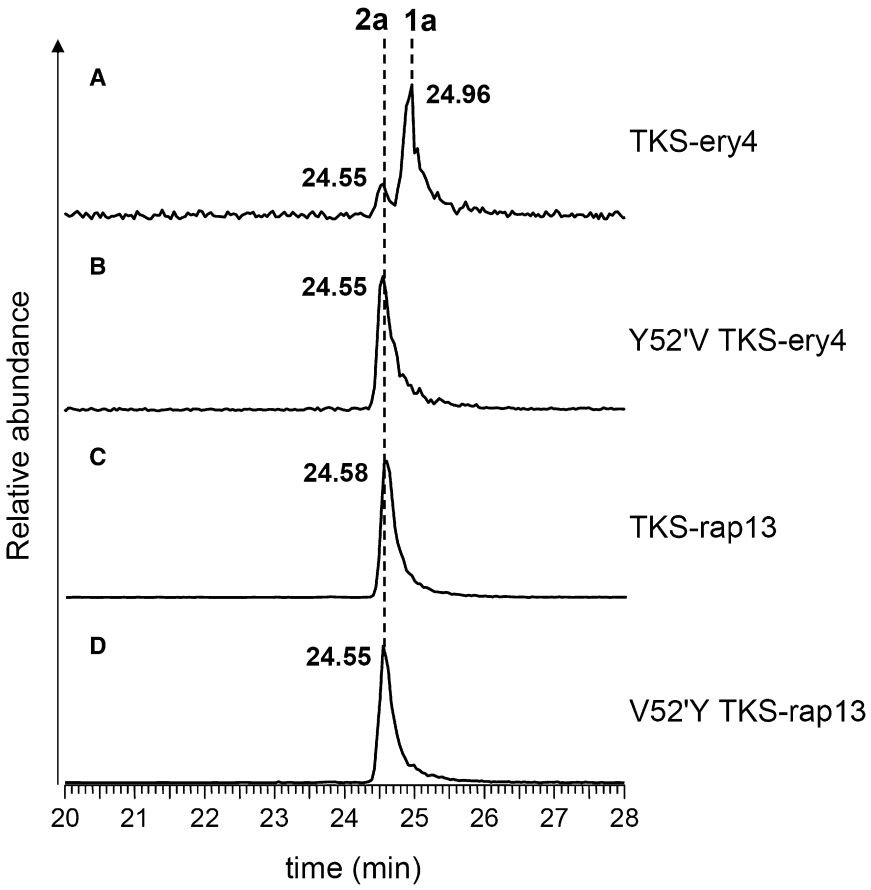

Figure 5. LC-MS Chromatograms of Triketide Lactones Produced from Engineered Triketide Synthases

(A–D) LC-MS traces of fermentation extracts from *S. erythraea* strains housing, respectively: (A) TKS-ery4; (B) Y52′V TKS-ery4; (C) TKS-rap13; and (D) V52′Y TKS-rap13.

and phthiocerol moieties are biosynthesized by PKS enzymes. All of the methyl branches of the DIMs found either in *Mycobacterium ulcerans*, the causative agent of the disease known as Buruli ulcer (Hong et al., 2008), or in the closely related *Mycobacterium marinum*, share the same configuration (Daffé et al., 1984). However, all are *opposite* in configuration to the methyl branches in the otherwise chemically identical DIMs found in tuberculosis-causing *M. tuberculosis* and in *Mycobacterium bovis* and *Mycobacterium leprae* (Figure 6) (Daffé and Lanéelle, 1988; Daffé, 1991).

A mycocerosic acid synthase (Mas) multienzyme has previously been identified in both *M. tuberculosis* H37Rv and *M. bovis* BCG, and shown to biosynthesize branched mycocerosic acids by the iterative elongation of *n*-fatty acyl-CoA with methylmalonyl-CoA followed by full reduction (Rainwater and Kolattukudy, 1983, 1985; Mathur and Kolattukudy, 1992). Scrutiny of sequence comparisons between the *M. tuberculosis* H37Rv and *M. bovis* BCG Mas proteins and their closest homologs in *M. ulcerans* Agy99 and *M. leprae* revealed that, in *M. ulcerans*, the ER domain of the Mas protein contains a Tyr52′ residue, corresponding to the all-*S* methyl-branched mycocerosate. In contrast, in *M. tuberculosis*, *M. bovis*, and *M. leprae*, which share the same all-*R* methyl-branch configuration in the mycocerosyl moieties of their DIMs, the 52′ residue of the Mas ER domain is phenylalanine (Supplemental Data).

For one mycobacterial PKS, the chirality of the product has not been determined with certainty. β-Mannosyl phosphomycoketides (MPMs) are important mycobacterial antigens produced in trace amounts by several pathogenic strains, including *M. tuberculosis*, *M. bovis*, and *M. avium* (Trivedi et al., 2005). MPMs are presented to T cells by CD1c of infected cells as part of the immune response (Moody et al., 2000; Matsunaga et al., 2004; Trivedi et al., 2005). The saturated fatty acid chains of MPMs exhibit an intriguing methyl-branching pattern (Figure 7). Originally thought to be of isoprenoid origin, MPMs are now known to be synthesized by the action of a specific PKS (Moody et al., 2000; Matsunaga et al., 2004), the product of *pks12*, the largest open reading frame in the *M. tuberculosis* genome. Pks12 is a bimodular PKS, which has been proposed to synthesize the fatty acyl chains of MPMs by elongating a long-chain acyl-CoA primer through five rounds each of alternating propionate and acetate units, derived, respectively, from methylmalonyl-CoA and malonyl-CoA (Figure 7) (Trivedi et al., 2005; Chopra et al., 2008). The first module of Pks12 is responsible for incorporating the methyl-branched units, and analysis of the sequence of the ER domain housed in this module showed that the 52′ position is always occupied by Phe and not Tyr (Supplemental Data). We therefore predict an *R* configuration for every methyl branch.

MPMs have not yet been isolated in amounts sufficient to determine their configuration directly. Meanwhile, others have undertaken the preparation of wholly synthetic MPMs containing all-*S* methyl branches, under the assumption (contrary to our present prediction) that this is the more likely configuration (van Summeren et al., 2006; de Jong et al., 2007). It is reported that the synthetic all-*S* MPM induces a T cell response comparable to that induced by an authentic MPM sample from *M. tuberculosis* (de Jong et al., 2007).

DISCUSSION

The factors governing stereocontrol during ER-catalyzed reduction on modular PKSs have been obscure. The MDR family of enzymes, to which the ER domains of modular PKS belong, is structurally heterogeneous, and crystallographic and mechanistic studies on enzymes most closely related to the ER domains have lagged behind work on, for example, the bacterial type II enoyl-(acyl carrier protein) reductases that are the target for

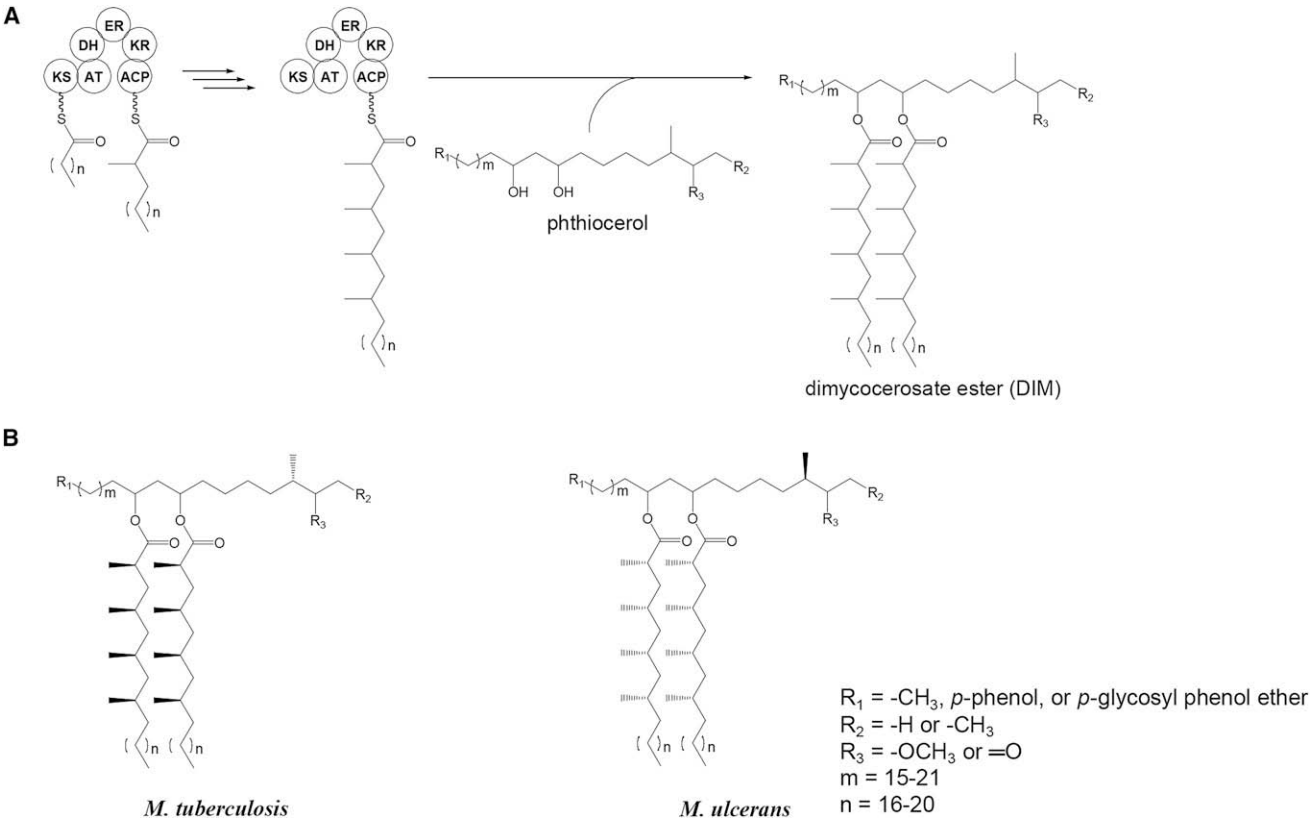

Figure 6. Biosynthesis of the Mycocerosyl Moiety of DIMs

(A) The Mas PKS catalyzes the biosynthesis of the mycocerosyl moieties of DIMS (Rainwater and Kolattukudy, 1985) by the iterative extension of *n*-fatty acyl chains by four successive propionate units, with ER-catalyzed enoylreduction occurring in every cycle.

(B) Configuration of DIMs in *M. tuberculosis* and *M. ulcerans*.

diazaborine antibiotics, the biocide triclosan, and one of the targets for the first-line antituberculosis drug isoniazid (Levy et al., 1999). In the latter enzymes, conserved lysine and tyrosine residues have been highlighted as important for catalysis, which are also important in the KR domains of PKS multienzymes (Reid et al., 2003). Meanwhile, stereocontrol on KR domains has been shown to involve subtle alterations in active site architecture that together determine the direction of approach of the substrate to a highly conserved catalytic apparatus (Baerga-Ortiz et al., 2006; O'Hare et al., 2006). We expected to uncover similar mechanisms at work in ER active sites, but nevertheless were surprised to make the observation that we report here, that a single amino acid residue differs systematically between PKS ER domains, yielding opposite configurations. It has previously been suggested that Tyr52 in *E. coli* QOR (and corresponding tyrosine residues in the homologous mammalian ζ-crystallins) might be essential for catalytic activity, acting, for example, to polarize the thioester carbonyl (Rao et al., 1997). However, the fact that many fully functional modular PKS ERs have valine at this position makes this proposal untenable. Our homology models suggest that another conserved tyrosine residue (Tyr125′ in either EryER4 or RapER13; Tyr130 of *E. coli* QOR) is a more plausible candidate for a catalytically essential residue, as previously proposed (Persson et al., 1994).

The results of specific mutations confirm that a tyrosine residue at 52′ is not essential for catalysis, and support a key role of the residue at this position in determining the stereochemical course of reduction. The replacement of this single residue in the EryER4 active site appears to be sufficient to alter the stereochemistry of enoylreduction (and without substantial diminution in overall yield of polyketide product) toward that seen with RapER13. In contrast to the switch of configuration previously achieved with a purified PKS KR domain (Baerga-Ortiz et al., 2006), the switch seen here occurs in the context of a PKS multienzyme, which encourages the view that highly targeted mutagenesis could become a useful tool in engineering PKSs to produce configurationally altered polyketide natural products. In the higher-yielding *S. erythraea* expression system, analysis showed that even the unmutated EryER4 domain gave rise to a small amount of triketide lactone **2a**, which has the "switched" (2*R*) configuration. Stereocontrol here may be imperfect, because the ER4 in its native context would be presented with a tetraketide rather than a diketide substrate. Possibly, too, the chain-terminating thioesterase (TE) domain, the natural heptaketide substrate of which has (2*R*) rather than (2*S*) configuration, might catalyze faster release of the triketide that leads to **2a** product. Taken together, the results suggest (as for the KR domains) that the stereochemical outcome reflects a delicate energetic balance between alternative ways in which the

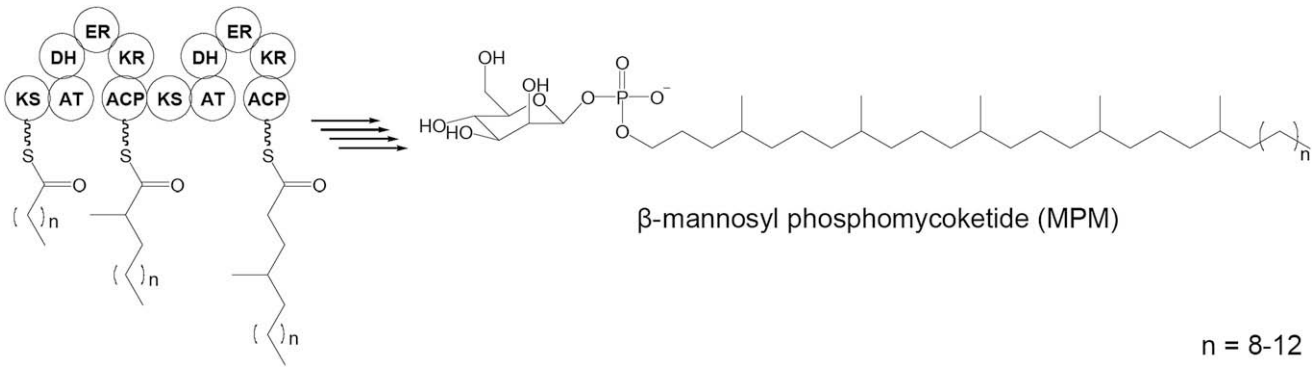

Figure 7. Biosynthesis of the Mycoketide Moiety of MPMs

The enzyme Pks12 catalyzes the biosynthesis of the mycoketide moiety of MPMs by iterative elongation of an *n*-fatty acyl chain with alternating propionate and acetate extender units (five cycles each) (Trivedi et al., 2005), with ER-catalyzed enoylreduction occurring in every cycle.

tethered substrate is presented to a highly conserved catalytic apparatus.

Site-directed mutagenesis (V52′Y) of the RapER13 domain gave an enzyme that showed no obvious impairment of catalytic function, and the triketide lactone product of which retained the same configuration as the unmutated enzyme. Resequencing of the appropriate region of DNA in the recombinant strain confirmed that the desired mutation had indeed been introduced. In addition, consistent results were obtained in each of the host strains *S. coelicolor* and *S. erythraea*. Further work is needed to establish the exact molecular basis for the resistance to the anticipated switch in the configuration of the product. These results suggest, however, that other residues at the active site remain to be identified that participate in stereocontrol, and they again highlight the finely balanced energetic differences between the different outcomes.

We have also shown here that the correlation of residue 52′ with the configuration of the methyl branches in fully reducing modules holds for certain mycobacterial modular PKS systems, such as Mas that synthesizes (multiply methyl-branched) mycocerosic acids. The ER domains of Mas can also be convincingly modeled on 1QOR as template. In contrast, other ER enzymes implicated in mycobacterial lipid formation appear to have a different fold. For example, a *trans*-acting ER, Rv2953, has been recently identified as catalyzing reduction of a double bond during biosynthesis of the phthiocerol moiety of DIM (Siméone et al., 2007). This gives rise to a methyl-branched stereocenter, which again in *M. ulcerans* is opposite in configuration to that in *M. tuberculosis*, *M. bovis*, and *M. leprae* (Daffé, 1991) (Figure 6). The active site residues participating in stereocontrol in Rv2953, and its homologs in other mycobacterial species, remain to be identified. For the multienzyme Pks12, which governs the production of MPMs in pathogenic strains of mycobacteria, our prediction is at odds with a previous proposal that was used to justify the synthesis of all-*S* rather than all-*R* MPM. However, given the potential importance of these lipids in modulating the host response to infection, it would be of considerable interest (pending the direct determination of the correct geometry of natural MPMs) to synthesize and determine the biological properties of the all-*R* methyl-branched MPMs as well. Meanwhile, identification of ER residue 52′ as tyrosine (or not) allows a robust prediction of methyl branch stereochemistry for the products of PKS multienzymes, the genes of which are uncovered through focused or genome-scale sequencing efforts.

SIGNIFICANCE

Exquisite stereochemical control is a central feature of the biosynthesis of polyketide antibiotics on modular polyketide synthases (PKSs). We have explored the molecular basis for the contribution to this control provided by the enoylreductase (ER) domain in those modules where a branching methyl group is introduced. By determining whether or not a unique tyrosine residue is present in the ER active site, the configuration of the polyketide product at the methyl branch can be confidently predicted. This further strengthens an emerging paradigm, according to which stereocontrol arises through the presentation of a different face of the substrate to an essentially invariant catalytic apparatus in each of the domains of a PKS. The finely balanced energetic difference between these presentational modes is highlighted by our finding that mutation of this key residue in an intact PKS system can switch the configuration of the product. This both illustrates a potential mechanism for natural evolution of these multienzymes, and encourages further experiments to alter that specificity by active site engineering.

EXPERIMENTAL PROCEDURES

Homology Modeling

The sequences of erythromycin PKS ER4 (EryER4) and rapamycin PKS ER13 (RapER13) were compared with all protein sequences deposited in Swiss Prot, using the Blast v2.2 program (Altschul et al., 1997). A PSI-BLAST search in each case produced an alignment between the EryER4 and RapER13 domains and homologous proteins. Homologous proteins with known structure were identified using the FUGUE homology recognition server (Shi et al., 2001). FUGUE searches for homologs in the structural profile library derived from the structure-based alignments in the HOMSTRAD database (Mizuguchi et al., 1998), and uses the environment-specific substitution tables to generate, automatically, the best alignments for the top hits. The homolog found with highest percentage identity (24%) to both ER domains was *E. coli* QOR (PDB ID 1QOR). The alignments produced by FUGUE for this highest-scoring hit were formatted and analyzed as previously described (Baerga-Ortiz et al., 2006). The models of EryER4 and RapER13 were constructed with MODELER

(Sali and Blundell, 1993) (using *E. coli* QOR as the template) and improved using loop-building programs, as previously described (Baerga-Ortiz et al., 2006). Final energy and structure minimization were performed using the SYBYL force field (Tripos, Inc., St. Louis, MO). The models were validated by PROCHECK (Laskowski et al., 1993) and by visual inspection using 3D graphics software.

Construction of Plasmids

Constructs for S. coelicolor *CH999*

For TKS-ery4, the plasmid pJLK41 (Kellenberger et al., 2008) was subjected to site-directed mutagenesis using primers 1 and 2 (Supplemental Data) to introduce the mutation Y52′V, resulting in plasmid pJLK41mut. The NdeI/XbaI fragment of this plasmid was cloned into vector pCJR133, digested with the same enzymes, to yield plasmid pYH141.

For TKS-raps13, the NdeI/XbaI fragment of pJLK27 (Kellenberger et al., 2008) was ligated to the vector pCJR133 (Rowe et al., 1998), digested with the same enzymes, yielding plasmid pYH136. The mutation V52′Y was introduced by site-directed mutagenesis using primers 3 and 4 (Supplemental Data), resulting in plasmid pYH137.

Constructs for S. erythraea BIOT1717-JC2

For TKS-ery4, the NdeI/XbaI fragment of pJLK41 was ligated into pFS1 digested with the same enzymes, resulting in plasmid pDK18.0. The NdeI/XbaI fragment of pJLK41mut was subcloned into pFS1 to yield pDK18.1.

For TKS-raps13, the NdeI/XbaI fragment of plasmid pJLK27 was inserted into vector pFS1 (comprising the ΦVWB integrase, an *attB* integration site, the DEBS native promoter, an origin of transfer for conjugation, and an apramycin resistance gene), digested with the same enzymes, resulting in plasmid pDK19.0. The variant TKS-raps13-V52′Y was likewise subcloned into pFS1, resulting in plasmid pDK19.1.

Site-Directed Mutagenesis

Site-directed mutagenesis was carried out using either the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene), following the manufacturer's instructions, or by overlap-extension PCR (Ho et al., 1989), using the primers mentioned above. The mutations were confirmed by sequencing (DNA Sequencing Facility, Department of Biochemistry, University of Cambridge, Cambridge, UK).

Manipulation and Handling of Bacterial Strains

E. coli was routinely grown in 2TY medium (1.6% tryptone, 1% yeast extract, 0.5% NaCl, containing 25 μg ml^{-1} kanamycin, 50 μg ml^{-1} apramycin, or 25 μg ml^{-1} chloramphenicol, if appropriate) at 37°C and transformed and maintained by standard procedures. *E. coli* ET12567/pUZ8002 (Kieser et al., 2000) was used for conjugation, and *E. coli* DH10B for general cloning purposes.

The plasmids pYH137 and pYH141 were introduced into *S. coelicolor* by conjugation (Kieser et al., 2000). The resulting strains were grown at 30°C in TSBY liquid medium (3% tryptone soy broth, 10.3% sucrose, 0.5% yeast extract) for strain maintenance and isolation of genomic DNA and on tap water medium agar (0.5% glucose, 1% sucrose, 0.5% tryptone, 0.25% yeast extract, 2% agar, pH 7.1, tap water) for the analysis of triketide lactone production.

Plasmids pDK18.0, pDK18.1, pDK19.0, and pDK19.1 were introduced into *S. erythraea* BIOT1717-JC2 by conjugation (Kieser et al., 2000). For strain maintenance and isolation of genomic DNA, M79 medium (1% glucose, 1% peptone, 0.2% yeast extract, 0.6% NaCl, 1% casein hydrolysate, pH 7.0, tap water) or TSB medium (Bacto BD Tryptic Soy Broth) was used. For analysis of triketide lactone production, eryP medium (5% glucose, 3% soybean flour, 0.3% $(NH_4)SO_4$, 0.5% NaCl, 0.6% $CaCO_3$, Antifoam A Concentrate [Sigma], 50 μg ml^{-1} apramycin, pH 7.0) was used, at 30°C.

Analysis of Triketide Lactones

To isolate triketide lactone from *S. coelicolor* CH999 strains, a tap water medium agar plate was extracted after 7 days of growth with ethyl acetate (1.6% formic acid). The solvent was removed in vacuo, and the residual oil dissolved in 600 μl methanol for analysis by LC-MS.

For the isolation from *S. erythraea* BIOT1717-JC2, a 10 ml TSB medium (50 μg ml^{-1} apramycin, 25 μg ml^{-1} nalidixic acid) was grown at 30°C for 3 days. A 1 ml aliquot of this starter culture was used to inoculate 25 ml of eryP medium (Rowe et al., 1998) (50 μg ml^{-1} apramycin). After 10 days at 30°C, triketide lactone was extracted with 25 ml ethyl acetate (containing 1.6% formic acid), the organic phase was removed in vacuo, and the remaining oil redissolved in 1 ml methanol before being subjected to LC-MS analysis.

LC-MS analysis was carried out on an Agilent 1200 HPLC system fitted with a Phenomenex ODS column (250 × 2.00 mm, 5 μm) and a Finnigan LTQ mass spectrometer system. The mobile phase used was a linear gradient from 95% water, 5% acetonitrile, 0.1% formic acid to 30% water, 70% acetonitrile, 0.1% formic acid over 35 min.

SUPPLEMENTAL DATA

Supplemental Data include two figures and two tables and can be found with this article online at http://www.chembiol.com/cgi/content/full/15/11/1231/DC1/.

ACKNOWLEDGMENTS

We thank T. Stinear and K.J. Weissman for helpful advice. This work was supported by a project grant from the Biotechnology and Biological Sciences Research Council (UK) to P.F.L., and by grants from the Cambridge Commonwealth Trust and the Natural Sciences and Engineering Research Council of Canada and a Universities UK—Overseas Research Scholarship to D.H.K. P.F.L. and S.F.H. designed the study. B.P. carried out homology modeling. S.F.H., Y.S., D.H.K., F.S., J.S. and H.H. carried out experiments. D.H.K. and P.F.L. wrote the manuscript. This article is dedicated to Manfred T. Reetz on the occasion of his 65th birthday.

Received: July 30, 2008
Revised: September 11, 2008
Accepted: September 17, 2008
Published: November 21, 2008

REFERENCES

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

Amy, C.M., Witkowski, A., Naggert, J., Williams, B., Randhawa, Z., and Smith, S. (1989). Molecular cloning and sequencing of cDNAs encoding the entire rat fatty acid synthase. Proc. Natl. Acad. Sci. USA *86*, 3114–3118.

Baerga-Ortiz, A., Popovic, B., Siskos, A.P., O'Hare, H.M., Spiteller, D., Williams, M.G., Campillo, N., Spencer, J.B., and Leadlay, P.F. (2006). Directed mutagenesis alters the stereochemistry of catalysis by isolated ketoreductase domains from the erythromycin polyketide synthase. Chem. Biol. *13*, 277–285.

Bedorf, N., Dietmar, S., Gerth, K., Reichenbach, H., and Höfle, G. (1993). Isolation and structure elucidation of soraphen $A_{1\alpha}$, a novel antifungal macrolide from *Sorangium cellulosum*. Liebigs Ann. Chem. *1993*, 1017–1021.

Bock, M., Buntin, K., Muller, R., and Kirschning, A. (2008). Stereochemical determination of thuggacins A-C, highly active antibiotics from the myxobacterium *Sorangium cellulosum*. Angew. Chem. Int. Ed. Engl. *47*, 2308–2311.

Caffrey, P. (2003). Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases. Chembiochem *4*, 649–662.

Camacho, L.R., Constant, P., Raynaud, C., Lanéelle, M.A., Triccas, J.A., Gicquel, B., Daffé, M., and Guilhot, C. (2001). Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*. J. Biol. Chem. *276*, 19845–19854.

Castonguay, R., He, W., Chen, A.Y., Khosla, C., and Cane, D.E. (2007). Stereospecificity of ketoreductase domains of the 6-deoxyerythronolide B synthase. J. Am. Chem. Soc. *129*, 13758–13769.

Celmer, W.D. (1965). Biogenetic, constitutional, and stereochemical unitary principles in macrolide antibiotics. Antimicrob. Agents Chemother. *5*, 144–156.

Chopra, T., Banerjee, S., Gupta, S., Yadav, G., Anand, S., Sirolia, A., Roy, R.P., Mohanty, D., and Gokhale, R.S. (2008). Novel intermolecular iterative mechanism for biosynthesis of mycoketide catalyzed by a bimodular polyketide synthase. PLoS Biol. *6*, e163. 10.1371/journal.pbio.0060163.

Cortés, J., Wiesmann, K.E.H., Roberts, G.A., Brown, M.J.B., Staunton, J., and Leadlay, P.F. (1995). Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage. Science *268*, 1487–1489.

Cox, J.S., Chen, B., McNeil, M., and Jacobs, W.R. (1999). Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice. Nature *402*, 79–83.

Daffé, M. (1991). Further stereochemical studies of phthiocerol and phenolphthiocerol in mycobacteria. Res. Microbiol. *142*, 405–410.

Daffé, M., and Lanéelle, M.A. (1988). Distribution of phthiocerol diester, phenolic mycosides and related compounds in mycobacteria. J. Gen. Microbiol. *134*, 2049–2055.

Daffé, M., Lanéelle, M.A., Roussel, J., and Asselineau, C. (1984). Lipides spécifiques de *Mycobacterium ulcerans*. Ann. Microbiol. (Paris) *135A*, 191–201.

de Jong, A., Arce, E.C., Cheng, T.Y., van Summeren, R.P., Feringa, B.L., Dudkin, V., Crich, D., Matsunaga, I., Minnaard, A.J., and Moody, D.B. (2007). CD1c presentation of synthetic glycolipid antigens with foreign alkyl branching motifs. Chem. Biol. *14*, 1232–1242.

Gokhale, R.S., Saxena, P., Chopra, T., and Mohanty, D. (2007). Versatile polyketide enzymatic machinery for the biosynthesis of complex mycobacterial lipids. Nat. Prod. Rep. *24*, 267–277.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using polymerase chain reaction. Gene *77*, 51–59.

Hong, H., Demangel, C., Pidot, S.J., Leadlay, P.F., and Stinear, T. (2008). Mycolactones: immunosuppressive and cytotoxic polyketides produced by aquatic mycobacteria. Nat. Prod. Rep. *25*, 447–454.

Janssen, D., Albert, D., Jansen, R., Mÿller, R., and Kalesse, M. (2007). Chivosazole A-elucidation of the absolute and relative configuration. Angew. Chem. Int. Ed. Engl. *46*, 4898–4901.

Keatinge-Clay, A.T. (2007). A tylosin ketoreductase reveals how chirality is determined in polyketides. Chem. Biol. *14*, 898–908.

Kellenberger, L., Galloway, I.S., Sauter, G., Böhm, G., Hanefeld, U., Cortés, J., Staunton, J., and Leadlay, P.F. (2008). A polylinker approach to reductive loop swaps in modular polyketide synthases. Chembiochem. *9*, 2740–2749.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., and Hopwood, D.A. (2000). Practical Streptomyces Genetics (Norwich, UK: The John Innes Foundation).

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Levy, C.W., Roujeinikova, A., Sedelnikova, S., Baker, P.J., Stuitje, A.R., Slabas, A.R., Rice, D.W., and Rafferty, J.B. (1999). Molecular basis of triclosan activity. Nature *398*, 383–384.

Ligon, J., Hill, S., Beck, J., Zirkle, R., Molnár, I., Zawodny, I., Money, S., and Schupp, T. (2002). Characterization of the biosynthetic gene cluster for the antifungal polyketide soraphen A from *Sorangium cellulosum* So ce26. Gene *285*, 257–267.

Maier, T., Jenni, S., and Ban, N. (2006). Architecture of mammalian fatty acid synthase at 4.5 Å resolution. Science *311*, 1258–1262.

Marsden, A.F., Caffrey, P., Aparicio, J.F., Loughran, M.S., Staunton, J., and Leadlay, P.F. (1994). Stereospecific acyl transfers on the erythromycin-producing polyketide synthase. Science *263*, 378–380.

Mathur, M., and Kolattukudy, P.E. (1992). Molecular cloning and sequencing of the gene for mycocerosic acid synthase, a novel fatty acid elongating multifunctional enzyme from *Mycobacterium tuberculosis* var. *bovis Bacillus Calmette-Guerin*. J. Biol. Chem. *267*, 19388–19395.

Matsunaga, I., Bhatt, A., Young, D.C., Cheng, T.Y., Eyles, S.J., Besra, G.S., Briken, V., Porcelli, S.A., Costello, C.E., Jacobs, W.R., Jr., and Moody, D.B. (2004). *Mycobacterium tuberculosis* pks12 produces a novel polyketide presented by CD1c to T cells. J. Exp. Med. *200*, 1559–1569.

McDaniel, R., Ebert-Khosla, S., Fu, H., Hopwood, D.A., and Khosla, C. (1994). Engineered biosynthesis of novel polyketides: influence of a downstream enzyme of the catalytic specificity of a minimal aromatic polyketide synthase. Proc. Natl. Acad. Sci. USA *91*, 11542–11546.

Mizuguchi, K., Deane, C.M., Blundell, T.L., and Overington, J.P. (1998). HOMSTRAD: a database of protein structure alignments for homologous families. Protein Sci. *7*, 2469–2471.

Moody, D.B., Ulrichs, T., Mühlecker, W., Young, D.C., Gurcha, S.S., Grant, E., Rosat, J.P., Brenner, M.B., Costello, C.E., Besra, G.S., et al. (2000). CD1c-mediated T-cell recognition of isoprenoid glycolipids in *Mycobacterium tuberculosis* infection. Nature *404*, 884–888.

O'Hare, H.M., Baerga-Ortiz, A., Popovic, B., Spencer, J.B., and Leadlay, P.F. (2006). High throughput mutagenesis of active site residues to evaluate models of stereochemical control in ketoreductase domains from the erythromycin-producing polyketide synthase. Chem. Biol. *13*, 287–296.

Persson, B., Zigler, J.S., and Jörnvall, H. (1994). A super-family of medium-chain dehydrogenase/reductases (MDR). Eur. J. Biochem. *226*, 15–22.

Rainwater, D.L., and Kolattukudy, P.E. (1983). Synthesis of mycocerosic acids from methylmalonyl coenzyme A by cell-free extracts of Mycobacterium tuberculosis var. bovis BCG. J. Biol. Chem. *258*, 2979–2985.

Rainwater, D.L., and Kolattukudy, P.E. (1985). Fatty acid biosynthesis in *Mycobacterium tuberculosis* var. *bovis Bacillus Calmette- Guérin*. Purification and characterization of a novel fatty acid synthase, mycocerosic acid synthase, which elongates *n*-fatty acyl-CoA with methylmalonyl-CoA. J. Biol. Chem. *260*, 616–623.

Rao, P.V., Gonzalez, P., Persson, B., Jörnvall, H., Garland, D., and Zigler, J.S. (1997). Guinea pig and bovine ζ-crystallins have distinct functional characteristics highlighting replacements in otherwise similar structures. Biochemistry *36*, 5353–5362.

Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D.V., Hutchinson, C.R., and McDaniel, R. (2003). A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. Biochemistry *42*, 72–79.

Rowe, C.J., Cortés, J., Gaisser, S., Staunton, J., and Leadlay, P.F. (1998). Construction of new vectors for high-level expression in actinomycetes. Gene *216*, 215–223.

Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. *234*, 779–815.

Shi, J., Blundell, T.L., and Mizuguchi, K. (2001). FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J. Mol. Biol. *310*, 243–257.

Siméone, R., Constant, P., Guilhot, C., Daffé, M., and Chalut, C. (2007). Identification of the missing *trans*-acting enoyl reductase required for phthiocerol dimycocerosate and phenolglycolipid biosynthesis in *Mycobacterium tuberculosis*. J. Bacteriol. *189*, 4597–4602.

Smith, S., and Tsai, S.-C. (2007). The type I fatty acid and polyketide synthases: a tale of two megasynthases. Nat. Prod. Rep. *24*, 1041–1072.

Staunton, J., and Weissman, K.J. (2001). Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. *18*, 380–416.

Thorn, J.M., Barton, J.D., Dixon, N.E., Ollis, D.L., and Edwards, K.J. (1995). Crystal structure of *Escherichia coli* QOR quinone oxidoreductase complexed with NADPH. J. Mol. Biol. *249*, 785–799.

Trivedi, O.A., Arora, P., Vats, A., Ansari, M.Z., Tickoo, R., Sridharan, V., Mohanty, D., and Gokhale, R.S. (2005). Dissecting the mechanism and assembly of a complex virulence mycobacterial lipid. Mol. Cell *17*, 631–643.

van Summeren, R.P., Moody, D.B., Feringa, B., and Minnaard, A.J. (2006). Total synthesis of enantiopure β-D-mannosyl phosphomycoketides from *Mycobacterium tuberculosis*. J. Am. Chem. Soc. *128*, 4546–4547.

Udwary, D.W., Zeigler, L., Asolkar, R.N., Singan, V., Lapidus, A., Fenical, W., Jensen, P.R., and Moore, B.S. (2007). Genome sequencing reveals complex

secondary metabolome in the marine actinomycete *Salinispora tropica*. Proc. Natl. Acad. Sci. USA *104*, 10376–10381.

Wenzel, S.C., Williamson, R.M., Grünanger, C., Xu, J., Gerth, K., Martinez, R.A., Moss, S.J., Carroll, B.J., Grond, S., Unkefer, C.J., et al. (2006). On the biosynthetic origin of methoxymalonyl-acyl carrier protein, the substrate for incorporation of "glycolate" units into ansamitocin and soraphen. J. Am. Chem. Soc. *128*, 14325–14336.

Wiesmann, K.E.H., Cortés, J., Brown, M.J.B., Cutter, A.L., Staunton, J., and Leadlay, P.F. (1995). Polyketide synthesis *in vitro* on a modular polyketide synthase. Chem. Biol. *2*, 583–589.

Weissman, K.J., Timoney, M., Bycroft, M., Grice, P., Hanefeld, U., Staunton, J., and Leadlay, P.F. (1997). The molecular basis of Celmer's rules: the stereochemistry of the condensation step in chain extension on the erythromycin polyketide synthase. Biochemistry *36*, 13849–13855.

